SETHU S
DATA SCIENTIST | DATA ANALYST

|*www.sethus4791@gmail.com* | *www.linkedin.com/in/sethus4791* | *www.kaggle.com/sethu123123* |
|*www.github.com/itsmesethus* | *PH No: +91 8122815260 | Address: Villupuram, Tamil Nadu - 605103|*

PROFESSIONAL SUMMARY

Dedicated Data Scientist and Analyst with a strong record of extracting actionable insights from complex data. Proficient in Statistical Analysis, Hypothesis Testing, and Data Modeling, utilizing Python, SQL, and R. Demonstrated leadership through project delivery and a commitment to innovation. Continuously upskilling in data-driven solutions, with a drive to elevate team success.

WORK EXPERIENCE

EVOSCIEN UK- STATISTICIAN   **BANGALORE, INDIA**
May 2023 to Present

- ✓ *Boosted data-driven decision accuracy by 30% through leading data acquisition, consolidating datasets, and applying advanced statistical methodologies.*
- ✓ *Pioneered Entomological Research direction by designing and executing innovative Hypothesis Tests, leading to the validation of critical project hypotheses.*
- ✓ *Optimized Experimental Design protocols in collaboration with Entomology teams, resulting in a 20% improvement in project outcomes and data accuracy.*
- ✓ *Transformed complex Entomological data into actionable patterns, creating compelling Data Visualizations, and streamlined reporting, increasing project efficiency by 25%.*

PROJECTS

MULTICLASS DISEASE CLASSIFICATION OF MEDICAL IMAGE DATA USING CONVOLUTIONAL NEURAL NETWORK --*M.sc Final year Project*   *(Jan 2023 – Mar 2023)*

- ✓ *Engineered and fine-tuned convolutional neural networks (CNNs) using TensorFlow and Keras, resulting in an exceptional accuracy of 94.88% for classifying Pneumonia X-ray images.*
- ✓ *Demonstrated mastery in multi-class classification, achieving remarkable accuracies: 95.67% for Potato Plant Leaf Disease (3-class), 95.18% for Corn Plant Leaf Disease (4-class), and 90.87% for Gastrointestinal Disease in Humans (8-class).*
- ✓ *Employed VGG19 with strategic layers, optimizing via Stochastic Gradient Descent to achieve superior 8-class classification performance.*

BCG DATA SCIENCE VIRTUAL INTERNSHIP - CHURN BUSTER: PREDICTIVE MODELING TO CUSTOMER RETENTION *-- Self Project*   *(Aug 2023 – Sep 2023)*

- ✓ *Orchestrated the deployment of diverse predictive models, encompassing ANN (Logistic Regression), CatBoost, XGBoost, GradientBoost, and KNeighbors Classifiers, to achieve an outstanding F1 score exceeding 95% with balanced accuracy.*
- ✓ *Strategically optimized data processes, delivering a remarkable 30% reduction in prediction errors through advanced techniques such as SMOTE and the Yeo-Johnson transformation, ensuring the readiness of predictive modeling.*
- ✓ *Translated machine learning insights into effective strategies, resulting in substantial gains in customer retention and cost savings. Additionally, crafted impactful Power BI dashboards for data visualization.*

COGNIZANT ARTIFICIAL INTELLIGENCE VIRTUAL INTERNSHIP – GALA GROCERIES: PRECISION STOCKING WITH ML -- *Self Project*   *(Sep 2023 – Oct 2023)*

- ✓ *Facilitated comprehensive EDA to uncover vital patterns and insights within the retail stock price dataset, providing a solid foundation for subsequent modeling and optimization.*

- ✓ *Leveraged Ensemble regressors & other models such as AdaBoost, Hist-Gradient-Boosting, Radius Neighbors, Bagging Regressors, and ANN-based Multiple Linear Regression to achieve over 94% R^2 precision in retail stock price predictions.*
- ✓ *Showcased exceptional model precision, achieving an extraordinary MSE of approximately 0.0523 and MAE of around 0.131, reaffirming the reliability of stock price forecasts.*

QUANTIUM DATA ANALYTICS VIRTUAL INTERNSHIP – RETAIL STORE PERFORMANCE ANALYSIS AND INSIGHTS -- self *Project*   *(May 2023 – Jun 2023)*

- ✓ *Transformed store data analysis for 272 locations, elevating accuracy through thorough data validation. Discovered a remarkable 15% sales surge via rigorous trial and control store assessments.*
- ✓ *Strategically segregated stores using advanced statistical methods, revealing a 12% performance differential. Employed Correlation Analysis and Two-Way ANOVA to identify key drivers.*
- ✓ *Equipped the Category Manager with a data-driven report, spurring strategic decisions. Guided actions resulted in a 10% sales boost and store performance optimization.*

EDUCATION

Bharathiar University   **Coimbatore, Tamil Nadu**
Master Of Science in Statistics- **CGPA: 8.61**   *Sept 2021 – May 2023*

- ✓ <u>*Relevant coursework:*</u> *Descriptive Statistics, Sampling Theory, Probability Theory, Statistical Estimation Theory, Statistical Quality Control, Multivariate Statistical Analysis, Econometrics, Statistical Inference, Programming in R, Design of Experiments and Stochastic Processes.*

Arignar Anna Governments Arts College   **Villupuram, Tamil Nadu**
Bachelor of Science in Statistics - **CGPA: 9.39**   *Jun 2018 – May 2021*

SKILLS & INTERESTS

- ✓ **PROGRAMMING LANGUAGES:** *Python, R and SQL.*
- ✓ **LIBRARIES:** *Pandas, Numpy, Matplotlib, Seaborn, Plotly, Scipy, Sckit-Learn, TensorFlow, Keras, Statsmodels, Imblearn , Pingouin etc.*
- ✓ **TOOLS:** *Power BI, IBM SPSS, MINITAB, STATISTICA, Microsoft Excel, MYSQL.*
- ✓ **DATA PREPROCESSING AND OTHERS:** *Data Cleaning, Data Analytics, Exploratory Data Analysis (EDA), Feature Engineering, Feature Selection Techniques, Data Visualization, Outlier Detection, Correlation Analysis, Statistical Tests etc.*
- ✓ **ML EXPERTISE (Supervised and Unsupervised ML):** *Regression, Classification, Clustering, Predictive Modelling, Quantitative Analysis, Statistical Modelling, CNN.*
- ✓ **SOFT SKILLS:** *Goal setting, Research Skills, Decision-making, Active listening, Optimism, Collaboration, Mind-mapping, Networking and Critical observation.*
- ✓ **INTERESTS:** *Machine Learning, Reading Blogs, Listening to music.*

CERTIFICATIONS & WORKSHOPS

- ✓ *IBM Data Science Professional Certificate – (IBM-Coursera, Jan 2023 – Jul 2023).*
- ✓ *Machine Learning Specialization (Supervised ML-Regression & Classification, Advanced Learning Algorithms, Unsupervised Learning & Recommender Systems)– (DeepLearning.Ai - Coursera, Oct 2023)*
- ✓ *Ensemble Methods In Python – (DataCamp, Sep 2023)*
- ✓ *Feature Engineering For Machine Learning In Python – (DataCamp, Sep 2023)*
- ✓ *Data Analytics With Python – (NPTEL, Swayam, Jan 2022- Apr 2022).*
- ✓ *Databases And SQL For Data Science With Python - (IBM - Coursera, Sep 2022 – Nov 2022).*
- ✓ *Statistical Sciences And Its Application Using Python – (RUSA 2.0 - BCTRC, Bharathiar University, Aug 2022).*
- ✓ *Data Visualization With Python – (IBM – Coursera, Jun 2023 – Jul 2023)*